

EFFICIENT SUBBAND ENCODING OF MAGNITUDE/PHASE SPECTRA

A. K. Soman and P. P. Vaidyanathan

Department of Electrical Engineering
California Institute of Technology
Pasadena, CA 91125

Abstract

Subband coding schemes have been used before to efficiently encode data. The theory of perfect-reconstruction systems has also been studied extensively. For certain types of signals however, perfect reconstruction is not essential. It is sufficient to have knowledge of either the magnitude or the phase spectrum only. For example, in encoding speech signals, phase distortion can be tolerated, and filter banks which introduce phase distortion have been used for encoding speech signals. However, these schemes have an inherent redundancy: the samples are still transmitted at the sampling rate. In this paper we will show how magnitude or phase information can be encoded at only half the sampling rate in a filter bank formalism.

1 Introduction

Subband coding has been employed in the past to encode speech and image signals. The theory of maximally decimated perfect reconstruction filter banks has also been studied extensively [1]-[3]. The main advantage of the subband coding formalism is that it allows us to treat different frequency regions differently. For example, in presence of quantization this can give rise to coding gain. Consider Fig. 1 which shows an example of a two-channel filter bank. The signal $\hat{X}(z)$ can be expressed as [3]

$$\begin{aligned}\hat{X}(z) &= 0.5[H_0(z)F_0(z) + H_1(z)F_1(z)]X(z) \\ &+ 0.5[H_0(-z)F_0(z) + H_1(-z)F_1(z)]X(-z) \quad (1) \\ &= T(z)X(z) + A(z)X(-z). \quad (2)\end{aligned}$$

The second term is the aliasing term, and can be cancelled by choosing $F_0(z) = H_1(-z)$ and $F_1(z) = -H_0(-z)$. Having cancelled the aliasing term, the transfer function of the system becomes $T(z) = 0.5[H_0(z)H_1(-z) - H_1(z)H_0(-z)]$. Perfect reconstruction is ensured by choosing the analysis filters $H_0(z)$ and $H_1(z)$ such that the transfer function $T(z) = cz^l$.

In several applications such as speech processing, perfect reconstruction is not essential. It is sufficient to know only the magnitude spectrum of the input $x(n)$. In other words, $T(z)$ need not be a delay, but is allowed to be an allpass function. This gives a greater freedom in designing the analysis and synthesis filters. Such filter banks have been studied for the two channel case in [7], and for the M channel case in [8]. The overall transfer function of the filter bank system is an allpass function, which guarantees that there is no amplitude distortion. Phase distortion however, is present.

On the other hand, information about the phase spectrum is important for image signals. In [9], the author has discussed how the original image can be obtained from the phase spectrum under certain conditions. For image signals, it would therefore be sufficient to ensure that the systems does not introduce phase distortion.

While the information regarding the magnitude or phase spectra can be extracted using the method outlined above, these schemes have an inherent redundancy. The overall transmission rate is equal to the rate of the input signal. In other words, we transmit data over M channels, at a rate r/M each, which gives an overall rate of r , equal to that of the input. We now ask: If one desires information about the magnitude spectra or phase spectra but not both, can it be done by transmitting samples below the rate of the input (Nyquist rate)? Moreover, can this be done in a filter bank formalism, so as to retain all the inherent advantages of subband coding?

Our work shows that the answers to the above questions are in the affirmative. The purpose of this paper is to show how magnitude or phase information can be recovered by transmitting samples at only half the Nyquist rate, using a digital filter bank.

This work was supported in parts by NSF grant MIP 8919196, Tektronix Inc. and Rockwell International.
0-7803-1254-6/93\$03.00 © 1993 IEEE

2 A Cepstrum based Approach to Encode Magnitude/Phase Spectra

The complex cepstrum of a signal $x(n)$ is defined as the signal $\hat{x}(n)$ [4] where

$$\hat{X}(e^{j\omega}) = \log|X(e^{j\omega})| + j\angle[X(e^{j\omega})]. \quad (3)$$

In this section we will achieve the stated goal through an approach based on computing the cepstrum of the input signal. Following are the steps in the scheme.

STEP 1. The input signals is first divided into blocks. In the case of speech, one normally desires the knowledge of short-time spectra and blocks of 20 ms (200 samples) are often used. On the other hand, images are inherently of finite length. We denote by $X(z)$ the z -transform of this block-input $x(n)$ of length L . Its complex cepstrum is $u_1(n)$. Since $x(n)$ is a real signal of finite length, its complex cepstrum is also real, but of infinite length. However, we have [4]:

Property 1: For a rational $X(z)$ with no zeros on the unit circle, the complex cepstrum decays at least as fast as $c/|n|$, where c is a constant, and n is the sample index.

Knowing the above, we can truncate the cepstrum by multiplying it with a symmetric window of length $2 * K + 1$, where we choose $2 * K + 1 \approx L$. Denote this truncated cepstrum as $u(n)$ and its z -transform as $U(z)$. Hence

$$U(e^{j\omega}) \approx \log|X(e^{j\omega})| + j\angle[X(e^{j\omega})]. \quad (4)$$

STEP 2. We then form the sequence $w(n)$ where

$$W(z) = U(z) + U(z^{-1}). \quad (5)$$

$W(z)$ therefore satisfies the property that $W(z^{-1}) = W(z)$, i.e., $w(n)$ is symmetric, and has odd length. Also, with the above construction, we have the relation

$$W(e^{j\omega}) = 2\text{Re}[U(e^{j\omega})] = 2\log|X(e^{j\omega})|. \quad (6)$$

STEP 3. The sequence $w(n)$ is then subband encoded through a digital filter bank. We choose the filter banks developed in [5],[6] to encode $w(n)$. The reason is that this filter bank has analysis filters that are both of odd length and symmetric. This means that the subband signals after downsampling are also symmetric. This in turn implies that only half the number of samples in each subband are required for reconstruction. These samples are independently quantized and

transmitted. There are other filter banks with linear phase filters. However, for our purpose it is essential that both the filters be of odd length. If the filters are not of odd length, it can be verified that the subband signal in that channel *after subsampling* will not be symmetric.

By using the filter bank in [5], [6], we have therefore transmitted data at only half the rate of the input sequence.

STEP 4. At the synthesis end, we reconstruct the signal $w(n)$, and compute its inverse cepstrum. In fact, it is sufficient to find the inverse logarithm of its Fourier transform. This gives us the approximate magnitude spectrum of the input signal $x(n)$.

Remarks: There are several points to be noted in the above scheme. 1) The signal transmitted is $w(n)$, and its Fourier transform is given by eq.(6). The logarithmic transformation is monotone, and hence the spectrum of $w(n)$ has different energies across subbands if the original signal $x(n)$ has such a property. Hence subband coding of $w(n)$ makes sense.

2) We note that the signal $w(n)$ is in fact symmetric. It would appear that it is sufficient to transmit only half of $w(n)$. However, transmitting only half of $w(n)$ would not retain its spectral characteristics, and subband coding would not make sense. Hence the filter bank with odd-length symmetric filters is crucial to the above scheme.

Cepstral techniques have been employed in the past to process speech signals [4]. Our scheme permits us to exploit the additional advantages of subband processing, and that too at half the rate. As an example, consider Fig. 2(a), which shows the magnitude spectrum of a predominantly lowpass signal. Fig. 2(b) shows the magnitude spectrum after having been encoded and recovered using our procedure. The bit rate used was 1.5 bits/sample.

In order to recover the phase information, we use $W(z) = U(z) - U(z^{-1})$ instead of eq. (5). Then, we would have

$$W(e^{j\omega}) = 2j\angle[U(e^{j\omega})] = 2j\angle[X(e^{j\omega})]. \quad (7)$$

Now, $w(n)$ is an antisymmetric signal, and the subband signals after quantization are also antisymmetric. Thus it is necessary to transmit only half of them.

3 A DFT Based Approach

In the previous section, we outlined a cepstrum based approach for encoding the phase spectrum. There we used Property 1 for truncating the cepstrum by multiplying it with a window. Therefore, the approximation in eq. (4) is good only when the block-length is large. If the length of the signal is small, or if we deal with blocks of the image at a time, one can employ DFT based methods to encode the phase/magnitude spectra at half the sampling rates. The reason for restricting the size of the input will be clarified subsequently. It has been shown in [9] that under specific conditions one can reconstruct images given only their phase spectrum. We will now demonstrate the method to reconstruct images after subband coding their phase.

The input image is first divided into blocks. We have chosen blocks of size 32 pixels by 32 pixels. We then compute the two dimensional DFT of each block. The magnitude spectrum shows a low-pass nature, where most of the energy is concentrated into a few coefficients. The phase spectrum on the other hand, does not have any such distribution. However, we still encode the phase with number of bits depending on the magnitude of the response at that frequency point. In this way, we encode the phase of only those DFT coefficients whose magnitude is significant. The rationale for such a bit-allocation is that the magnitude of the DFT coefficient decides its relative importance. Hence the phase corresponding to DFT coefficients with highest magnitude needs to be represented with largest number of bits. Since the phase response is antisymmetric, we can transmit only half the coefficients.

There are two ways to perform the reconstruction. The first method [9] is based on solving a system of equations. It has been shown in [10] that a sequence $x(n_1, n_2)$ can be retrieved from its phase (upto a scale factor) if the signal is real, of finite length, and its Fourier Transform cannot be factored into lower order polynomials in $e^{j\omega_1}$ and $e^{j\omega_2}$. We know that the phase response satisfies eq. (1.43) in [9]. If (ω_1, ω_2) are distinct points, that gives rise to N^2 equations in N^2 unknowns. These are solved to give the original image. We performed experiments to solve for the image after subband quantizing the phase. However, quantization is found to substantially degrade the solution obtained by this method.

The other method is based on modelling the magnitude response of the block as an AR(1) process as has been done in [9]. The subband quantized phase is combined with the AR(1) magnitude to obtain the recon-

structed image. Modelling of the image as an AR(1) process restricts the size of the image, and therefore we have treated only blocks of the image at a time. Moreover, it means that we transmit only two additional coefficients per block. Figs. 3(a) shows the original block (an eye) from the image of Lena, and Fig. 3(b) shows the reconstructed image. In order to observe the effect of phase quantization on sharp edges, we performed the same experiment on Fig. 4(a). The reconstruction is shown in Fig. 4(b). The bit rate used was 0.42 bpp in both cases.

References

- [1] M. J. T. Smith and T. P. Barnwell, III, "A Procedure for Designing Exact Reconstruction Filter Banks for Tree-Structured Sub-band Coders," Proc. of IEEE Int. Conf. on ASSP, San Diego, March 1984.
- [2] M. Vetterli, "A Theory of Multirate Filter Banks," IEEE Trans. on ASSP, pp.356-372, Mar. 1987.
- [3] P. P. Vaidyanathan, "Multirate Digital Filters, Filter Banks, Polyphase Networks, and Applications: A Tutorial," Proc. of IEEE, Vol.78, No.1, Jan. 1990.
- [4] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*, Prentice Hall, 1989.
- [5] M. Vetterli and D. Le Gall, "Analysis and Design of Perfect Reconstruction Filter Banks Satisfying Symmetry Constraints," Proc. 22nd Conf. Info. Sci. Syst., Princeton, NJ, Mar. 1988.
- [6] T. Nguyen and P. P. Vaidyanathan, "Two Channel Perfect Reconstruction FIR QMF Structures Which Yield Linear-Phase Analysis Filters," IEEE Trans. on ASSP, pp. 676-690, May, 1989.
- [7] P. P. Vaidyanathan, P. Regalia, S. K. Mitra, "Design of Doubly Complementary IIR Filter using a Single Complex Allpass Filter with Multirate Applications," IEEE Tr. Cir. Sys., Apr. 1987.
- [8] M. Renfors, T. Saramaki, "Recursive N the Band Digital Filters, Parts I and II," IEEE Tr. Cir. Sys., Jan 1987.
- [9] J. Lim, *Two-dimensional Signal and Image Processing* Prentice Hall, 1990.
- [10] M. Hayes, "The Reconstruction of a Multidimensional Sequence from the Phase or Magnitude of its Fourier Transform," IEEE Trans. ASSP, Apr. 1982.

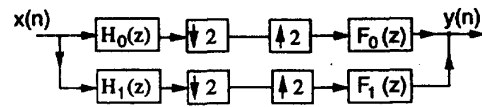


Fig. 1. A 2-channel Filter bank

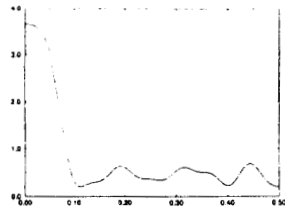


Fig. 2(a). Original Magnitude Spectrum

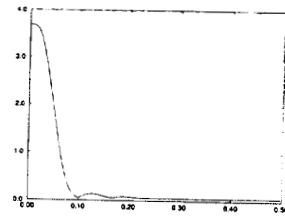


Fig. 2(b). Reconstructed Magnitude Spectrum

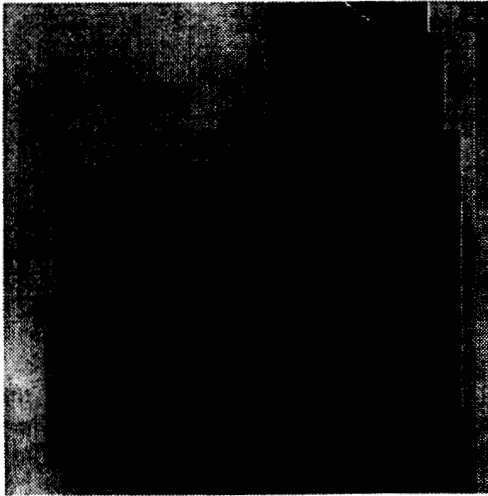


Fig. 3(a). Original Block from image

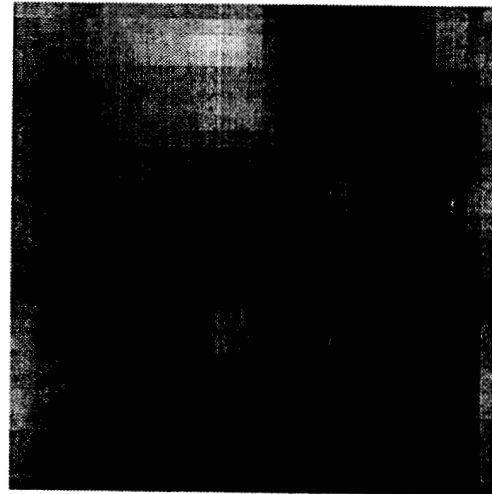


Fig. 3(b). Reconstructed block

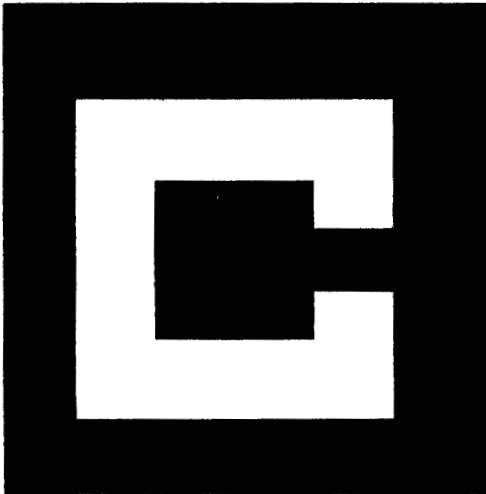


Fig. 4(a). Original Alphabet

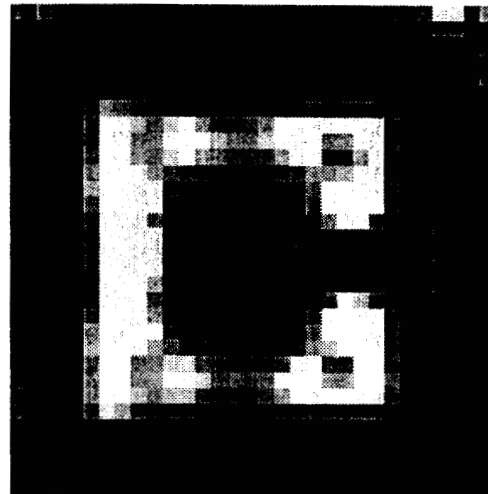


Fig. 4(b). Reconstructed Alphabet